

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:32 ; Search time 8498.8 Seconds

(without alignments)
31.610 Million cell updates/sec

Title: US-09-851-670-19

Perfect score: 25

Sequence: 1 gctgcgtcgtgacccctcttgc 25

Scoring table: IDENTITY_NUC

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 111874

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gD_estl:*
11: gD_estl2:*
12: gD_hic:*
13: gD_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17	68.0	42	13	AZ938933 2M0197N10
C 2	14.6	58.4	43	13	A1115359 un84c06.r
C 3	14.2	56.8	30	13	A2310102 1M001BD24
C 4	14.2	56.8	32	13	BE317690 NF053E10L
C 5	14	56.0	55	11	BE977884 bs70d09.y
C 6	13.8	55.2	50	10	AU103198 AU103198
C 7	13.6	54.4	27	10	AM059815 LE7f08.y9
C 8	13.6	54.4	52	11	R42396 yf92a07.si
C 9	13.6	54.4	57	10	A1019839 ua91h02.r
C 10	13.4	53.6	51	13	B05502 cSR1-64f6-u
C 11	13.4	53.6	60	10	AM004453 701931490
C 12	13.2	52.8	56	13	AZ820546 2M0092207

13	13.2	52.8	58	10	AA415880
14	13.2	52.8	60	13	A2575892
15	13	52.0	32	13	A2306035
C 16	13	52.0	38	13	A2471063
C 17	13	52.0	39	13	A2428906
C 18	13	52.0	39	13	A2788235
C 19	13	52.0	42	11	T17509
C 20	12.8	51.2	36	13	A2807406
C 21	12.8	51.2	41	13	A2576441
C 22	12.8	51.2	52	11	BE643936
C 23	12.8	51.2	53	10	A1906734
C 24	12.8	51.2	55	13	A2755909
C 25	12.8	51.2	58	11	D38739
C 26	12.6	50.4	34	10	AA933070
C 27	12.6	50.4	46	10	A1357720
C 28	12.6	50.4	47	13	A2814250
C 29	12.6	50.4	51	10	AA672337
C 30	12.6	50.4	56	10	A1310538
C 31	12.6	50.4	58	10	A1340421
C 32	12.6	50.4	58	10	AA696109
C 33	12.6	50.4	60	11	BE857801
C 34	12.4	49.6	34	13	A2775556
C 35	12.4	49.6	39	13	A2610841
C 36	12.4	49.6	47	10	AA658865
C 37	12.4	49.6	49	10	A1567998
C 38	12.4	49.6	50	10	A1010205
C 39	12.4	49.6	50	10	A10107283
C 40	12.4	49.6	52	11	BE651215
C 41	12.4	49.6	54	13	A2612285
C 42	12.4	49.6	55	10	A1522415
C 43	12.4	49.6	60	11	BE638573
C 44	12.2	48.8	29	13	A0073365
C 45	12.2	48.8	32	10	A0014466

ALIGNMENTS

RESULT 1
AZ938933 42 bp DNA GSS 26-APR-2001
LOCUS 2M0197N10R Mouse 10kb plasmid UNGC2M library Mus musculus genomic
DEFINITION clone UNGC2M0197N10 R, DNA sequence.

ACCESSION AZ938933.1 GI:13799132

VERSION GSS.

KEYWORDS house mouse.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 42)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinger,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0197 row: N column: 10

Seq primer: CACACGAGAAACGCTATGACC
Class: plasmid ends

High quality sequence stop: 42.

FEATURES
Source Location/Qualifiers
1..42

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U0GC2M0197N10"
/clone_1lb="Mouse 10kb plasmid U0GC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notice="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g11473211419b1Af129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

Query Match	68.0%	Score 17	DB 13	Length 42
Best Local Similarity	80.0%	Pred. No. 2.5e+03		
Matches	20	Conservative	0	Mismatches 5; Indels 0; Gaps 0;
Oy	1	gctgcagctctgagatccctctcttgc	25	
Db	41	gctagaaactctgagatcccatcttggc	17	

RESULT	2
A1115359	
LOCUS	A1115359 43 bp mRNA EST 02-SEP-1998
DEFINITION	uh84c06.r1 Soares mouse urogenital ridge NMR mus musculus cDNA clone IMAGE:1764394 5' similar to SW-6472_RAT P53678 CLATHRIN COAT ASSEMBLY PROTEIN APA7 HOMOLOG 2 ;, mRNA sequence.

ACCESSION	AI115359	GI:3515683
VERSION	AI115359.1	
KEYWORDS	EST	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

Eucaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.

TITLE	The Washu-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Maira M/Mouse EST Project

Email: mousestewartson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MCI:963918
Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 1.
Location/Qualifiers

1..43

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="1764394"
/clone_lib="Soares mouse urogenital ridge NKUR"
/sex="equal ratio of male:female"
/tissue_type="urogenital ridge (embryonic)"
/dev_stage="fetal, mixture of 11.5 and 12.5 dpc"
/lab_host="DH10B"

/note="Organ: gonad; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGAAGTCGAGCGCGCGATTTCTTTTTTTTTTTTTTTTTTTT T3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

5 a 11 c 7 g 14 t 6 others

BASE COUNT
ORIGIN

Query Match	58.48%	Score 14.6:	DB 10:	Length 43:
Best Local Similarity	77.38%	Pred. No. 2.5e+04:		
Matches 17:	Conservative	0:	Mismatches 5:	Indels 0:
				Gaps 0:
0y	1	gctgcatctgfatcccttctt	22	
db	12	gctctnntgctgcatcttctt	33	

RESULT	3
AZ310102	
LOCUS	AZ310102 30 bp DNA GSS 29-SEP-2000
DEFINITION	I00018D2R Mouse 10k plasmid UGCGIM library Mus musculus genomic
clone	U06C1M0018D24 R, DNA sequence.
AZ310102	
ACCESSION	AZ310102.1 GI:10351755
VERSION	

SOURCE ORGANISM	REFERENCE AUTHORS
house mouse. <i>Mus musculus</i> .	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus (1 (pages 1 to 30))	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C., ...

TITLE	Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL	plasmid inserts
COMMENT	Unpublished (2000)
	Contact: Robert B. Weiss

8412, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0018 row: D column: 24
Seq primer: CACACAGCAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 30.
location/Qualifiers
1. .30

```
1.30
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="UUCGIM0018D24"
/clone_1lb="Mouse 10kb plasmid UUCGIM library"
```

/sex="Male"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pMD42ny, Purified genomic DNA from M. musculus C57Bl/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b/AP129072.1), a copy number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
2 a 14 c 5 g 9 t

Query Match 56.8%; Score 14.2; DB 13; Length 30;
Best Local Similarity 84.2%; Pred. No. 3.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ctgcactgtatccctc 20
||||| ||||| |||||
Db 12 CTCGCTGTGTCAGCCCTTC 30

RESULT 4
LOCUS BE317690 52 bp mRNA EST 21-DEC-2000
DEFINITION NF053E10LF1071 Developing leaf Medicago truncatula cDNA clone
ACCESSION BE317690
VERSION BE317690
KEYWORDS BE317690.2 GI:11960827
SOURCE EST.
ORGANISM Medicago truncatula
barrel medic.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eucolecyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 52)
AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.
TITLE Expresed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula leaf library
JOURNAL Unpublished (2000)
COMMENT On Jul 14, 2000 this sequence version replaced g1:9191563.
CONTACT: May GD

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Medicago Genome Initiative accession: MGI:S:21873
Insert Length: 776 Std Error: 0.00
Plate: 053 row: E column: 10
Seq primer: TCACACGAGAAACAGCTATGAC.
Location/Qualifiers

FEATURES
source
1..52
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF053E10LF"
/clone_id="Developing leaf"
/tissue_type="leaf"

/dev_stage="pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very young, developing, mature and senescing leaves."

BASE COUNT
ORIGIN
11 a 16 c 14 g 11 t

Query Match 56.8%; Score 14.2; DB 10; Length 52;
Best Local Similarity 84.2%; Pred. No. 3.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ctgcactgtatccctc 20
||||| ||||| |||||
Db 38 CTCGAGGTGTCATACCTTC 20

RESULT 5
LOCUS BE977884 55 bp mRNA EST 04-OCT-2000
DEFINITION bs70d09.y1 Drosophila melanogaster adult testis library Drosophila melanogaster cDNA clone bs70d09 5', mRNA sequence.

ACCESSION BE977884
VERSION BE977884
KEYWORDS BE977884.1 GI:10608804
SOURCE EST.
ORGANISM Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 55)
AUTHORS Andrews, J., Bouffard, G. and Oliver, B.
TITLE Drosophila melanogaster testis expressed sequence tags
JOURNAL Unpublished (1999)
COMMENT Contact: Brian Oliver
Laboratory of Cellular and Developmental Biology
NIDDK, National Institutes of Health
6 Center Drive MSC 2715, Bldg 6, Rm B1-13, Bethesda, MD 20892 USA
Fax: (301) 496 5239
Email: oliver@helix.nih.gov,
<http://www.nidck.nih.gov/intram/people/boliver.htm>
Tissue isolation and library construction performed at the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (see <http://www.nidck.nih.gov/intram/people/boliver.htm>). DNA sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC; see <http://www.nisc.nih.gov>).
Plate: 70 row: d column: 09
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers

FEATURES
source
1..55
/organism="Drosophila melanogaster"
/strain="y[*] w[67c1] y"
/db_xref="taxon:7227"
/clone="bs70d09"
/clone_id="Drosophila melanogaster adult testis library"
/sex="male"
/dev_stage="1-5 day adult"
/lab_host="SOPK (Stratagene)"
/note="Organ: testis; Vector: pBluescript SK (Stratagene); Site_1: EcoR I; Site_2: Xho I; Testes dissected from 1-5 day adult y[*] w[67c1] Y males raised at 25°C. RNA isolated using Trizol (Life Technologies) and a single round of poly(A)+ selection using Oligotex (Qiagen). cDNA library constructed using Stratagene Zap-cDNA synthesis kit. Oligo dt-primed, size fractionated -1-6 kb, and directionally cloned at EcoRI and XhoI in uni-zap XR. Following a single round of amplification pBluescript SK phagemids were mass excised. A distribution channel for clones is being sought, but not currently available.
Requests for clones cannot be honored."

BASE COUNT
ORIGIN
8 a 11 c 19 g 17 t

Query Match 56.0%; Score 14; DB 11; Length 55;
 Best Local Similarity 77.3%; Pred. No. 4.5e+04;
 Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 cgaatctgataccctctcttc 25
 ||| | | | | | | | | |
 Db 4 CGAGCGATCATGCTCTCTTGC 25

RESULT 6
 AUI03198 50 bp mRNA EST 05-APR-2001
 LOCUS AUI03198 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION COLF5014, mRNA sequence.
 ACCESSION AUI03198
 VERSION AUI03198.1 GI:13552719
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 'H., Oca,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 'K., Suyama,A. and Sugano,S.
 Fine structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)
 JOURNAL Contact: Yutaka Suzuki
 COMMENT Department of Medical Science, University of Tokyo
 Institute of Medical Science, Minatoku, Tokyo 108-8639, Japan
 4-6-1, Shirokane-dai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 'S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
 source
 1..50
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="COLF5014"
 /clone_11b="Sugano Homo sapiens cDNA library"
 17 c 18 g 11 t

BASE COUNT 4 a 17 c 18 g 11 t

Query Match 55.2%; Score 13.8; DB 10; Length 50;
 Best Local Similarity 72.0%; Pred. No. 5.4e+04;
 Matches 18; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 gctgatactgataccctctcttc 25
 | | | | | | | | | |
 Db 20 GATGATCTGTCCTCTCTCGGC 44

RESULT 7
 AUI059815 27 bp mRNA EST 23-AUG-2000
 LOCUS AUI059815
 DEFINITION LE7108.yg DMC15 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AUI059815
 VERSION AUI059815.1 GI:6652137
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Brenner,S., Williams,S.R., Vermass,E.H., Storck,T., Moon,K.,
 McCallum,C., Mao,J.I., Kirchner,J.D., Elert,S., Dubridge,R.B.,
 Burcham,T. and Albrecht,G.
 In vitro cloning of complex mixtures of DNA on microbeads: Physical
 separation of differentially expressed cDNAs

TITLE
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)

MEDLINE 20144098

COMMENT

Contact: Burcham TS
 LYNX Therapeutics, Inc.
 25861 Industrial Blvd., Hayward, CA 94545, USA
 Tel: 510 670 9338
 Fax: 510 670 9302
 Email: tim@lynxgen.com
 Sequence obtained from LYNX Therapeutics Megasort technology.
 Collected from the down-regulated gate.
 High quality sequence stop: 27.

FEATURES
 source

1..27
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="DNC15"
 /cell_type="monocytic leukemia"
 /cell_line="THP-1 (THP-202)"
 /note="Vector: PCR2.1; Cloning of PCR products from
 micro-beads carrying 3' end of down-regulated cDNA. THP-1
 cells non-induced (treated with DMSO only)."
 3 a 9 c 7 g 8 t

BASE COUNT 3 a 9 c 7 g 8 t

Query Match 54.4%; Score 13.6; DB 10; Length 27;
 Best Local Similarity 80.0%; Pred. No. 6e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 atctgataccctctcttc 25
 | | | | | | | | | |
 Db 8 ACCTGATCTCTCTCTCTGC 27

RESULT 8
 R42396 52 bp mRNA EST 22-MAY-1995
 LOCUS R42396/c
 DEFINITION yf92a07.s1 Soares Infant brain INTB Homo sapiens cDNA clone
 IMAGE:29952 3' similar to gb|879041HUMAN|NE531 Human carcinoma
 cell-derived Alu RNA transcript, (RNA): gb:567859 TRANSCRIPTION
 INITIATION FACTOR IIE-ALPHA CHAIN (HUMAN);, mRNA sequence.
 ACCESSION R42396
 VERSION R42396.1 GI:800620
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
 'M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisan,E., Waterston
 'R., Williamson,A., Wohlmann,P. and Wilson,R.
 The Washu-Merck EST Project
 Unpublished (1995)
 JOURNAL Contact: Wilson RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Insert Size: 1268
 High quality sequence starts: 1 High quality sequence stops: 1
 Source: IMAGE Consortium, LNM. This clone is available royalty-free
 through LNM; contact the IMAGE Consortium (info@image.lln.gov)
 for further information. Trace considered overall poor quality
 Insert Length: 1268 Std Error: 0.00
 Seq Primer: Promega -21m13
 High quality sequence stop: 1.

FEATURES
 source
 1..52
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="GDB:40229"
 /db_xref="taxon:9606"

Query Match	Best Local Similarity	54.4%	Score 13.6;	DB 11;	Length 52;
Matches 16;	Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0;
0y	5 gatcgtgatccctcttgc 25				
Db	37 GANCTGACATCAGTCATTGC 17				
RESULT 9					
LOCUS	A1019839				
DEFINITION	u91h02.r1 Soares_mammary_gland_NbMNG Mus musculus cDNA clone IMAGE:1364883 5' similar to gp:f00516 Mouse mRNA for aldolase A (MOUSE);, mRNA sequence.				
ACCESSION	A1019839				
VERSION	A1019839.1				
KEYWORDS	GI:3234175				
SOURCE	EST.				
ORGANISM	house mouse.				
	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 57)				
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Stepien,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
TITLE	The WashU-HMHI Mouse EST Project				
JOURNAL	Unpublished (1996)				
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMHI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:898103 Seq primer: -28ml3 rev2 ET from Amersham.				
FEATURES					
SOURCE	Location/Ouallifiers 1..57 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:1364883" /clone_lib="Soares_mammary_gland_NbMNG" /sex="male" /tissue_type="mammary gland" /dev_stage="4 weeks" /lab_host="DH10B" /note="Organ: mammary gland; Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - o1igo(dT)				

```

primer [5',
TGTTCACCAATCTGAGTGCGAGCGCCGCCGATGCTTTTTTTTTTTTTTTTTTTTTT
T 3'] : double-stranded cDNA was ligated to Eco RI
adaptors (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT      18 a       17 c       8 g       14 t
ORIGIN

Query Match          54.4% ; Score 13.6; DB 10; Length 57;
Best Local Similarity 80.0% ; Pred. No. 6,6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 gctcgatcgtgaccttc 20
        ||| ||||| |||||||
db      6 GCTAATCTGCATCCTTC 25

RESULT 10
BOS502   LOCUS           51 bp     DNA                GSS              13-JUL-1996
DEFINITION CSRJ-64f6-u CSRJ flow sorted Chromosome 11 specific cosmid Homo
sapientis genomic clone CSRJ-64f6, DNA sequence.
ACCESSION B05502
VERSION    B05502.1 GI:1414780
KEYWORDS   GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 51)
AUTHORS     Evans,G.A., Burpee,D., Davies,C., Hahner,L., Oliver,T., Gilbert,M.,
            Jones,D., Ward,T., Gillilan,E., Schagemann,J., Probst,S., Harris
            J., Derford,J., McFarland,J., Butzinski,K., Khan,M., Kupfer,K. and
            Garner,H.R.
            Genomic Sequence Sampled Map of Chromosome 11
            Unpublished (1996)
            Contact: Evans GA, Shane Growth
            McCormick Center for Human Growth and Development
            University of Texas Southwestern Medical Center At Dallas
            5323 Harry Hines Blvd, Dallas TX 75235-8591
            Tel: 214-648-1600
            Fax: 214-648-1666
            Email: gevas@uts.w.swned.edu, shane@mccormick.utswmed.edu
            Seq primer: T7
            Class: cosmid ends
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source           1..51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CSRJ-64f6"
                     /clone.lib="CSRJ flow sorted Chromosome 11 specific
                     cosmid"
                     /sex="female"
                     /cell.type="chimeric hamster somatic cell hybrid"
                     /note="Vector: sCOS-1; Human chromosome 11 specific cosmid
                     library prepared from flow sorted human Chromosome 11
                     derived from Chinese Hamster Ovary (CHO) monochromosomal
                     somatic cell hybrid, J1"
BASE COUNT      2 a       11 c       16 g       21 t       1 others
ORIGIN

Query Match          53.6% ; Score 13.4; DB 13; Length 51;
Best Local Similarity 70.8% ; Pred. No. 7,9e+04;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 gctcgatcgtgaccttc 24
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db	10	GNTAACTCTTGTCGCCCTTCTGTG	33
RESULT	11		
LOCUS	AM004453	60 bp	mRNA EST
DEFINITION	701931490 A. thaliana, mixed source Arabidopsis thaliana cDNA clone		08-SEP-1999
ACCESSION	701931490.		mRNA sequence.
VERSION	AM004453.1	GI:5851482	
SOURCE	EST.		
ORGANISM	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
REFERENCE	Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D., Wang,X., Hillman,J., Guejfer,K., Kim,C., Doyle,M., Brzoska,P., Gorogone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carplo,T., Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and Hanson.D. Arabiadopsids thailana Gene Expression Microarray Unpublished (1999)		
TITLE	Contact: David Smoller, Ph.D. Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc. 4633 World Parkway Circle, St. Louis, MO 63134, USA Tel.: 877-577-2733 Fax: 314-627-3334 Email: service@genomesystems.com.		
JOURNAL COMMENT	Location/Dualifiers 1..60 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="701931490" /clone_lib="A. thaliana, mixed source" /note-"This sequence was obtained from a clone generated with a PCR product of the target gene."		
FEATURES	BASE COUNT 5 a 19 c 10 g 24 t 2 others		
SOURCE	ORIGIN		
Query Match	Best Local Similarity	53.6%	Score 13.4;
Matches	17: Conservative	0:	Mismatches 7; Indels 0; Gaps 0;
Oy	ctcgatcgtgatccctcttgc	25	
Db		27	CITGNTCCCGTGGCTTTCTCTGC 50
RESULT	12		
LOCUS	AZ820546/c		
DEFINITION	2M0092J07R Mouse 10kb plasmid UUGCJM library Mus musculus genomic		
ACCESSION	AZ820546		
VERSION	AZ820546.1	GI:12990370	
KEYWORDS	GSS.		
SOURCE	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ORGANISM	Dunn,D., Ayragl,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000)		

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 306, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0092 row J column: 07
Seq primer: CACACAGGAACACTATGACC
Class: plasmid ends
High quality sequence stop: 56.
Location/Qualifiers

FEATURES

Source

1..56

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGM0092J07"
/clone_lib="Mouse 10kb plasmid UUCGM Library"
/sex="male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
<http://www.jax.org/resources/documents/dnares/>. The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (914732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT

24 a 8 c 19 g 5 t

ORIGIN

Query Match 52.8%; Score 13.2; DB 13; Length 56;
Best Local Similarity 83.3%; Pred. No. 9.6e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 4 cgactcgatcccttctc 21
|||||||
Db 20 CTACCTGTGATCCCTGT 3

RESULT 13
AA415880

LOCUS AA415880 58 bp mRNA EST 16-OCT-1997

DEFINITION v232903.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone

IMAGE:793396 5', mRNA sequence.

ACCESSION AA415880

VERSION AA415880.1 GI:2076148

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 58)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

The WashU-RHMI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:477804
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
source

Location/Qualifiers
1. 58
/organism="Mus musculus"
/strain="B6D2 F1/J"
/db_xref="taxon:10090"
/clone_id="IMAGE:793396"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pBluescribe (modified);
Site_1: Mu1; Site_2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dT): 5'-CGGTCGACCGTCGACGCTTTTCTTTT-3', CDNAS
were cloned into the Mu1/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 Kb."

BASE COUNT
ORIGIN

19 a 10 c 27 t

Query Match 52.8%; Score 13.2; DB 10; Length 58;
Best Local Similarity 83.3%; Pred. No. 9.6e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 atcgtgacccctctt 23
||||| ||||| |||||
Db 25 ATCTCTATCCCTCTTT 42

RESULT 14
A2575892 60 bp DNA GSS 06-DEC-2000
LOCUS AST-T23B0019 Genetrap T47D Human Breast Carcinoma Library Homo
DEFINITION sapiens genomic 5', DNA sequence.
VERSION A2575892
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 60)
Henkel,G., Livanage,M., Pratt,E., Huang,D., Riley,M., Bernardino,A.
Durrick,K. and Pollok,B.
Exon-trap tags from a T47D Genomescreen(TM) Library
Unpublished (2000)
Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Norreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com

Pools of cells were isolated from a Genomescreen(TM) library. The
library of cells was generated by retroviral integration of a gene
tagging element consisting of: 1) A promoterless beta-lactamase
preceded by a splice acceptor as a reporter for gene expression;
2) A promoter driving neomycin resistance followed by a splice
donor to trap downstream exons. 3' RACE from neomycin gene was
performed using total RNA from isolated pools. Output was shotgun

FEATURES
source

cloned in pAMP-1 and used to transform DH5-alpha competent
bacteria. 5' ends of reported sequences were immediately preceded
by splice donor from the trapping construct.
Class: exon-trapped.
Location/Qualifiers
1. 60

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/cell_line="T47D"
/note="Organ: Breast; Vector: pAMP-1; 3' RACE of total RNA
from genetrap pools; shotgun clone in pAMP-1 and used to
transform DH5-alpha competent bacteria."

BASE COUNT
ORIGIN

14 a 16 c 16 g 14 t

Query Match 52.8%; Score 13.2; DB 13; Length 60;
Best Local Similarity 83.3%; Pred. No. 9.7e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ctgcgtgacccctt 19
||||| ||||| |||||
Db 8 CTCGATCTGTCACCAT 25

RESULT 15
A2306035 32 bp DNA GSS 29-SEP-2000
LOCUS 1M0006M23R Mouse 10kb plasmid UUC1M library Mus musculus genomic
DEFINITION clone UUC1M0006M23 R, DNA sequence.
VERSION A2306035
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 32)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0006 row: M column: 23
Seq primer: CACACAGAAACAGCATGTGACC
Class: plasmid ends
High quality sequence stop: 32.
Location/Qualifiers
1. 32

FEATURES
source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="UUC1M0006M23"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adsorbed DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (q147321419b1dNF12072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT	6 a	12 c	3 g	11 t
ORIGIN				

ORIGIN

Query Match

52.08; Score 13; DB 13; Length 32;

Best Local Similarity	76.28;	Pred. No.	1.1e+05;
Matches	16;	Conservative	0;
		Mismatches	5;
		Indels	0;
		Gaps	0;

Matches	16;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY 2 ctcgatctgtgatcccttctt 22

Db 12 CCCCATCTGTAAATGCCCTCTT 32

```
Search completed: March  9, 2002, 00:09:35
Job time: 11051 sec
```

Job time: 11051 sec